JOURNAL OF COMBINATORIAL THEORY, Series A 40, 108–124 (1985)

General Forbidden Configuration Theorems

R. P. ANSTEE*

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada

Communicated by the Managing Editors

Received March 8, 1983

Results in this paper gives bounds on the number of columns in a matrix when certain submatrices are forbidden. Let F be a k by l (0, 1)-matrix with no repeated columns, column sums at least s. Let A be a m by n (0, 1)-matrix with no repeated column, column sums at least s and no submatrix F nor any row and column permutation of F. Then $n \leqslant \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{s}$. This bound is best possible for numerous F. The bound, with $s = 0$, is an easy corollary to a bound of Sauer and Perles and Shelah. The bounds can be extended to any F and to any F where we do not allow row and column permutations. The results follow from a configuration theorem that says, in essence, that matrices without a configuration are determined by row intersections of sets of rows of various sizes. A linear independence argument yields the bound. Results of Ryser, Frankl and Pach, Quinn and the author are obtained. © 1985 Academic Press, Inc.

1. INTRODUCTION

This paper presents some powerful forbidden configuration theorems which yield a general forbidden configuration theorem as a corollary. We show that if A is a (0, 1)-matrix on m rows with no repeated columns and no k by l configuration F, then the number of columns of A is bounded by a polynomial in m whose degree is roughly k. The bounds are not expected to be best possible in all cases but they are best possible in numerous special cases.

We concern ourselves exclusively with (0, 1)-matrices (hypergraphs if you prefer) and so we may forget to specify this from time to time. Usually in combinatorics, arbitrary row and column permutations leave the combinatorial properties of a matrix unchanged. Define a *configuration* as an equivalence class of matrices where one matrix represents another if it is a row and column permutation of the other matrix. A configuration is given by any representative. A matrix A is said to contain a configuration C if A

* Support provided by NSERC.

0097-3165/85 $3.00

has a submatrix B which is a representative of C. This notation is Ryser's [15]. We are interested in what happens when you forbid a configuration.

A property of (0, 1)-matrices is called *hereditary* if when A satisfies the property, then any submatrix, under any row and column permutation, satisfies the property. Thus if we consider the class of matrices satisfying a hereditary property, then a matrix B not satisfying the property yields a forbidden configuration for matrices in the class. Conversely, a property defined by forbidden configurations is hereditary.

Some very interesting examples exist. A matrix is *balanced*, as defined by Berge [6], if it has no configuration in the infinite set $\{C_k | k \text{ odd}, k \geqslant 3\}$, where C_k is given by a matrix of order k:

$$\begin{bmatrix} 1 & & & & 1\\ 1 & 1 & & 0\text{'s} & \\ & 1 & 1 & & \\ & & 1 & & \\ & 0\text{'s} & & \ddots & \\ & & & 1 & \\ & & & 1 & 1 \end{bmatrix}. \tag{1.1}$$

Balanced matrices have many interesting properties involving linear programming. *Totally balanced* matrices, much studied recently (e.g., Lovász [12], Anstee and Farber [4], Brouwer and Kolen [7]), can be defined by the set $\{C_k | k \geqslant 3\}$ of forbidden configurations. Matrices which are conformal (see Berge [6]) and are hereditary with this property can be defined by a single forbidden configuration C_3 [3]. Our results are for a single forbidden configuration but of course apply in general to give bounds on the number of columns in terms of the number of rows.

Another interesting problem involving forbidden configurations is Turan's problem. For an introduction see [9]. Define K_k^l to be the k by $\binom{k}{l}$ (0, 1)-matrix consisting of all possible columns of column sum l on k rows. The notation is reminiscent of K_m for the complete graph where row K_m^2 would be the incidence matrix and K_m^{m-1} the adjacency matrix. The complementary form of Turan's problem asks for a bound on n where A is an m by n (0, 1)-matrix with no repeated columns and no configuration K_k^l, where A has all column sums equal to l. We are unable to shed much light on these problems because of the restriction on column sums but the problem should help to motivate the reader. Another application of forbidden configurations can be found in the work of Bruen and Silverman [8], again for fixed column sums, who obtains results for finite geometries.

Our main result, in Section 3, is to show that if F is a configuration of size k by l with no repeated columns, column sums at least s, then an m by

n (0, 1)-matrix A with no repeated columns, column sums at least s, no configurations F, satisfies the bound $n \leqslant \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{s}$. This bound generalizes the bound of Sauer [18], Perles and Shelah [19], which considers the case $s=0$, $F=[K_k^k K_k^{k-1} \cdots K_k^0]$. Ryser [14] proved the bound for $s=2$, $F=K_3^2$ and Quinn [13] proved the bound for $s=k-1$, $F=K_k^{k-1}$. The case that F has repeated columns results in a weaker bound, $O(n^k)$. In any event, the number of columns is polynomially bounded, a surprising result.

In Section 5, we demonstrate that the same results yield polynomial bounds when we just forbid a submatrix F, not a configuration F. Our bounds are $n \leqslant \binom{m}{s-1} + \binom{m}{s-2} + \cdots + \binom{m}{0}$ for $s=kl$ and $s=13k \log_2 l$. Some results for special F, where the bound is not met, are discussed.

Section 2 is devoted to the basic configuration theorem which is quite powerful and general. The proof is based on an inductive argument of Ryser [15] later used in [2]. The theorem leads to precise bounds for certain configurations in Section 3. Here we use a generalization of Ryser's matrix equation for finite sets [17]. In Section 4 we discuss properties of the matrices meeting the bound and demonstrate their existence in certain cases. Füredi and Quinn's results [11], provide the most interesting examples.

2. The Forbidden Configuration Theorem

In this section we prove a quite general forbidden configuration theorem. The notation K_k^l, presented in the introduction, will be vital. We also need a concept that for the time being appears to have little relevance to the main result of the paper. Consider an $m \times n$ (0, 1)-matrix A. Let the *k-fold row intersection vector* of rows $\{r_1, r_2, \ldots, r_k\}$ be a row vector of length n with a 1 in column i if all the rows $r_1, r_2, \ldots, r_k$ have a 1 in column i, i.e., the Hadamard product of rows $r_1, r_2, \ldots, r_k$. This notation is due to Quinn [13]. We require all the indices $r_1, r_2, \ldots, r_k$ to be distinct. The 0-fold intersection vector will be the row vector of length n of all 1s. This useful notation is fairly natural. Define $A^{(l)}$ to be a set of entries indexed by all possible l-subsets of $\{1, 2, \ldots, m\}$, where the entries are the number of 1s in the l-fold row intersection vector given by the l rows indexed by the l-subset. Note that $A^{(0)}=m$ and $A^{(1)}$ and $A^{(2)}$ define the diagonal and off diagonal entries, respectively, of AA^{T}. Ryser proved the following results.

Theorem 2.1 [13]. *Let A, B be* (0, 1)*-matrices with no configuration K_3^2 and all column sums at least* 2. *Assume $A^{(2)}=B^{(2)}$. Then A and B are the same apart from a column permutation.*

THEOREM 2.2 [12]. *Let A, B be $(0, 1)$-matrices with A having no configuration $[K_3^2 K_3^0]$ and $[K_3^3 K_3^1]$. Assume $A^{(0)} = B^{(0)}$ (no. of columns equal), $A^{(1)} = B^{(1)}$, $A^{(2)} = B^{(2)}$. Then A and B are the same apart from a column permutation.*

We generalize these and other results in what follows. The proof uses the ideas of Ryser [15].

THEOREM 2.3. *Let s, t be given with $0 \leqslant s \leqslant t$. Let A, B be $(0, 1)$-matrices with column sums at least s and no configuration*

$$[K_t^t K_t^{t-2} \cdots K_t^s] \qquad \text{if} \quad s \equiv t \pmod 2,$$

or

$$[K_t^{t-1} K_t^{t-3} \cdots K_t^s] \qquad \text{if} \quad s \equiv t-1 \pmod 2.$$

Assume $A^{(s)} = B^{(s)}$, $A^{(s+1)} = B^{(s+1)}, \dots, A^{(t-1)} = B^{(t-1)}$. Then A and B are the same apart from a column permutation.

Proof. We prove this by induction on the number of rows, say m, of A and B. The result is trivially true for $m = 1, 2, \dots, t-1$. Consider matrices A, B satisfying the hypotheses where each is on m rows and $m \geqslant t$. Assume A and B have no matching columns. If they did, we could delete them in pairs and obtain smaller matrices satisfying the hypotheses. We need to show that A and B have no columns.

The proof involves showing that A and B should have a complicated structure using an inductive argument. Remember that we have to produce the forbidden configuration to get contradictions. The argument requires some difficult notation. Let $P_k^{\geqslant l}$ consist of all possible columns of column sum at least l on k rows where each individual column α may be repeated some number, λ_α, times where $\lambda_\alpha \in \{0, 1, 2, \dots\}$. The number λ_α is determined during the course of the proof. We write

$$\begin{bmatrix} P_k^{\geqslant l} \\ C(h) \end{bmatrix}, \tag{2.1}$$

to denote the following matrix. $C(h)$ stands for two matrices $C_0(h)$, $C_1(h)$. Under the λ_α copies of column α of $P_k^{\geqslant l}$, we have either $C_0(h)$ if the number of 1s of α is even or $C_1(h)$ if the number of 1s of α is odd. Thus the number of columns of $C_0(h)$ and $C_1(h)$ determine the λ_αs. We will be using the notation even if the matrices $C_0(h)$, $C_1(h)$ have no rows, just some number of columns (only for $h = 0$) or no columns, just some number of rows. Be patient, this does streamline the proof. Here is an example for $P_3^{\geqslant 1}$, where $C_0(h)$ has 3 columns and $C_1(h)$ has 2 columns.

$$\begin{bmatrix} P_3^{\geqslant 1} \\ C(h) \end{bmatrix}$$

$$= \left[\begin{array}{cc|cc|cc|ccc|ccc|ccc|cc} 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ \multicolumn{2}{c|}{C_1(h)} & \multicolumn{2}{c|}{C_1(h)} & \multicolumn{2}{c|}{C_1(h)} & \multicolumn{3}{c|}{C_0(h)} & \multicolumn{3}{c|}{C_0(h)} & \multicolumn{3}{c|}{C_0(h)} & \multicolumn{2}{c}{C_1(h)} \end{array}\right]. \tag{2.2}$$

We also need the notation

$$\begin{bmatrix} P_3^{\geqslant l} \\ \bar{C}(h) \end{bmatrix}, \tag{2.3}$$

where here $C_0(h)$ appears under copies of α if the number of 1s in α is odd and $C_1(h)$ appears when the number of 1s is even. The parity is reversed.

The index h in in $C(h)$ denotes that $C_i(h)$ has column sums h if $h<s$ or column sums at least s if $h=s$. Thus $C_i(0)$ is a zero matrix.

We wish to show that we may write

$$A = \begin{bmatrix} P_k^{\geqslant 0} & P_k^{\geqslant 1} & \cdots & P_k^{\geqslant s-1} & P_k^{\geqslant s} \\ C(s) & C(s-1) & & C(1) & C(0) \end{bmatrix},$$
$$B = \begin{bmatrix} P_k^{\geqslant 0} & P_k^{\geqslant 1} & \cdots & P_k^{\geqslant s-1} & P_k^{\geqslant s} \\ \bar{C}(s) & \bar{C}(s-1) & & \bar{C}(1) & \bar{C}(0) \end{bmatrix}, \tag{2.4}$$

where $C_0(l)$, $C_1(l)$ have no matching columns $0 \leqslant l \leqslant s$, for $k=0, 1,..., t$.

This claim is surely satisfied for $k=0$ in which case $C_0(s)=A$, $C_1(s)=B$. Assume the result is true for k as in (2.4). Separate off the $(k+1)$th row in the following way

$$\begin{bmatrix} P_k^{\geqslant 0} \\ C(s) \end{bmatrix} \rightarrow \begin{bmatrix} P_k^{\geqslant 0} & P_k^{\geqslant 0} \\ \delta & 1 \\ D(s) & E(s-1) \end{bmatrix}, \tag{2.5}$$

where δ stands for two possible row vectors δ_0, δ_1 occurring under λ_α copies of column α, where the parities match as in the definition of (2.1). The columns of $C_i(s)$, after the first row is deleted, divide into $D_i(s)$ and $E_i(s-1)$ depending on the column sum. Those columns yielding $E_i(s-1)$ must have a 1 in the first row of $C_i(s)$. We regroup the columns to obtain (2.5). For the remaining cases we have

$$\begin{bmatrix} P_k^{\geqslant l} \\ C(s-l) \end{bmatrix} \to \begin{bmatrix} P_k^{\geqslant l} & P_k^{\geqslant l} \\ 0 & 1 \\ D(s-l) & E(s-l-1) \end{bmatrix}, \qquad (0<l<s),$$

$$\begin{bmatrix} P_k^{\geqslant s} \\ C(0) \end{bmatrix} \to \begin{bmatrix} P_k^{\geqslant s} \\ 0 \\ D(0) \end{bmatrix}. \tag{2.6}$$

This yields, after a column permutation,

$$A = \begin{bmatrix} P_k^{\geqslant 0} & P_k^{\geqslant 0} & P_k^{\geqslant 1} & P_k^{\geqslant 1} & & P^{\geqslant s-1} & P_k^{\geqslant s-1} & P_k^{\geqslant s} \\ \delta & 1 & 0 & 1 & \cdots & 0 & 1 & 0 \\ D(s) & E(s-1) & D(s-1) & E(s-2) & & D(1) & E(0) & D(0) \end{bmatrix},$$

$$\tag{2.7}$$

$$B = \begin{bmatrix} P_k^{\geqslant 0} & P_k^{\geqslant 0} & P_k^{\geqslant 1} & P_k^{\geqslant 1} & & P_k^{\geqslant s-1} & P_k^{\geqslant s-1} & P_k^{\geqslant s} \\ \delta & 1 & 0 & 1 & \cdots & 0 & 1 & 0 \\ \bar{D}(s) & \bar{E}(s-1) & \bar{D}(s-1) & \bar{E}(s-2) & & \bar{D}(1) & \bar{E}(0) & \bar{D}(0) \end{bmatrix}.$$

We now delete the $(k+1)$th row and apply induction: the resulting matrices are equal apart from a column permutation when the columns of column sum less than s are deleted.

Checking the columns of column sum at least s with zeros in the first k rows, we deduce that $D_0(s)$ and $D_1(s)$ are the same apart from a column permutation. This works trivially if $D_0(0)$ and $D_1(s)$ have no columns.

Check the columns of column sum at least s with zeros in the first k rows except for a 1 in the row k. We deduce that $[D_1(s)\, E_1(s-1)\, D_1(s-1)]$ and $[D_0(s)\, E_0(s-1)\, D_0(s-1)]$ are the same apart from a column permutation. By hypothesis, $C_0(l)$ and $C_1(l)$ have no matching columns, $0 \leqslant l \leqslant s$, and so $D_0(s-1)$ and $D_1(s-1)$ have no matching columns. Thus we deduce that $E_1(s-1)$ and $D_0(s-1)$ are the same apart from a column permutation and that $E_0(s-1)$ and $D_1(s-1)$ are the same apart from a column permutation.

Inductively, check the columns of column sum at least s with zeros in the first k rows except for $s-l$ 1s in rows $k, k-1, \ldots, k-s+l+1$. We may deduce by the above argument that $E_1(l)$ and $D_0(l)$ are the same apart from a column permutation and that $E_0(l)$ and $D_1(l)$ are the same apart from a column permutation. This argument works for $s-l \leqslant k$. We note that $P_k^{\geqslant s-l}$ is empty for $s-l>k$ and so the result follows trivially for these cases. Note that the arguments hold even in the case that the matrices $D_i(0)$ and $E_i(0)$ have no rows, just some number of columns, which will occur for $m=t$ and $k=t-1$. We deduce that certain λ_αs are equal and hence we can set the pairs $D_0(0)$, $E_1(0)$ and $D_1(0)$, $E_0(0)$ to have equal numbers of columns even though they have no rows.

We now reorder the columns to obtain A and B as in (2.4) with k replaced by $k+1$. For $0 < l \leqslant s$, we have, by reordering columns,

$$\begin{bmatrix} P_k^{\geqslant l-1} & P_k^{\geqslant l} \\ 1 & 0 \\ E(s-l) & D(s-l) \end{bmatrix} \to \begin{bmatrix} P_{k+1}^{\geqslant l} \\ D(s-l) \end{bmatrix}, \tag{2.8}$$

noting that all possible columns of column sum l on $k+1$ rows are generated and using the deductions involving $E_1(s-l)$, $D_0(s-l)$ and $E_0(s-l)$, $D_1(s-l)$. For $l=0$, we note that $D_0(s)=D_1(s)$ forces δ_0 and δ_1 to not have any matching entries, i.e., $\delta_0+\delta_1=(1, 1,..., 1)$, otherwise we get matching columns in $C_0(s)$ and $C_1(s)$. Let $F_0(s)$ be that part of $D_0(s)$ lying under 0s of δ_0 and let $F_1(s)$ be that part of $D_0(s)$ lying under 1s of δ_0. Then, by reordening columns we may obtain

$$\begin{bmatrix} P_k^{\geqslant 0} \\ \delta \\ D(s) \end{bmatrix} \to \begin{bmatrix} P_k^{\geqslant 0} & P_k^{\geqslant 0} \\ 0 & 1 \\ F(s) & \bar{F}(s) \end{bmatrix} \to \begin{bmatrix} P_{k+1}^{\geqslant 0} \\ F(s) \end{bmatrix}. \tag{2.9}$$

Thus we may write after column permutation,

$$\begin{aligned} A &= \begin{bmatrix} P_{k+1}^{\geqslant 0} & P_{k+1}^{\geqslant 1} & \cdots & P_{k+1}^{\geqslant s-1} & P_{k+1}^{\geqslant s} \\ F(s) & D(s-1) & & D(1) & D(0) \end{bmatrix}, \\ B &= \begin{bmatrix} P_{k+1}^{\geqslant 0} & P_{k+1}^{\geqslant 1} & \cdots & P_{k+1}^{\geqslant s-1} & P_{k+1}^{\geqslant s} \\ \bar{F}(s) & \bar{D}(s-1) & & \bar{D}(1) & \bar{D}(0) \end{bmatrix}. \end{aligned} \tag{2.10}$$

This verifies our inductive claim that we may write A and B as in (2.4) for $k=0, 1,..., t$. Since $m \geqslant t$, it is possible for $m=t$ that the matrices $C_i(0)$ have no rows, just columns. For $k=t$ in (2.4) we have

$$\begin{aligned} A &= \begin{bmatrix} P_t^{\geqslant 0} & P_t^{\geqslant 1} & \cdots & P_t^{\geqslant s-1} & P_t^{\geqslant s} \\ C(s) & C(s-1) & & C(1) & C(0) \end{bmatrix}, \\ B &= \begin{bmatrix} P_t^{\geqslant 0} & P_t^{\geqslant 1} & \cdots & P_t^{\geqslant s-1} & P_t^{\geqslant s} \\ \bar{C}(s) & \bar{C}(s-1) & & \bar{C}(1) & \bar{C}(0) \end{bmatrix}. \end{aligned} \tag{2.11}$$

Assume $t \equiv s \equiv 0 \pmod 2$. Then if $C_0(l)$ has any columns at all for any l $(0 \leqslant l \leqslant s)$, then A has a configuration of all columns of even column sum in $P_t^{\geqslant s-l}$ and so the forbidden configuration $[K_t^t K_t^{t-2} \cdots K_t^s]$. If $C_1(l)$ has any columns at all, then B has a configuration consisting of all columns of even column sum in $P_t^{\geqslant s-l}$ and so the forbidden configuration. Thus A and B have no columns and the result holds.

The remaining three cases for the different parities of s and t yield the same conclusion. ▮

The proof extends to prove the generalization of Ryser's result,

THEOREM 2.4. *Let s, t be given with $0 \leqslant s \leqslant t$. Let A, B be $(0, 1)$-matrices with column sums at least s and with A having no configurations $[K_t^t K_t^{t-2} \cdots K_t^s]$ and $[K_t^{t-1} K_t^{t-3} \cdots K_t^{s+1}]$ for $s \equiv t \pmod 2$, or $[K_t^{t-1} K_t^{t-3} \cdots K_t^s]$ and $[K_t^t K_t^{t-2} \cdots K_t^{s+1}]$ for $s \equiv t-1 \pmod 2$. Then A and B are the same apart from a column permutation.*

Proof. Assume A and B have m rows. The result follows easily for $m = 1, 2, \ldots, t-1$. The same induction argument will apply and we will arrive at (2.11). Assume $t \equiv s \equiv 0 \pmod 2$. Then if $C_0(l)$ has any columns at all for any l $(0 \leqslant l \leqslant s)$, then A has a configuration of all columns of even column sum in $P_k^{\geqslant s-l}$ and so the forbidden configuration $[K_t^t K_t^{t-2} \cdots K_t^s]$. If $C_1(l)$ has any columns, we find that A has the configuration $[K_t^{t-1} K_t^{t-3} \cdots K_t^{s+1}]$. Thus A and B both have no columns as desired. ▮

The following result implies both Theorems 2.3 and 2.4 and is vital in Section 3.

PROPOSITION 2.5. *Let s, t be given with $0 \leqslant s \leqslant t$. Let A, B be nonnull $(0, 1)$-matrices with all column sums at least s, no matching columns and with $A^{(s)} = B^{(s)}$, $A^{(s+1)} = B^{(s+1)}, \ldots, A^{(t-1)} = B^{(t-1)}$. Then in some choice of t rows, one of A or B has the configuration $[K_t^t K_t^{t-2} \cdots K_t^s]$ for $s \equiv t \pmod 2$ or $[K_t^{t-1} K_t^{t-3} \cdots K_t^s]$ for $s \equiv t-1 \pmod 2$ and the other matrix has the configuration $[K_t^{t-1} K_t^{t-3} \cdots K_t^{s+1}]$ for $s \equiv t \pmod 2$ or $[K_t^t K_t^{t-2} \cdots K_t^{s+1}]$ for $s \equiv t-1 \pmod 2$.*

Proof. Prove the contrapositive, that if A and B do not have the forbidden configurations as specified in any choice of t rows then A, B are null. The same inductive argument applies to obtain (2.11). Assume $s \equiv t \equiv 0 \pmod 2$. Then if $C_0(l)$ has any columns, then A has the configuration $[K_t^t K_t^{t-2} \cdots K_t^s]$ and B has the configuration $[K_t^{t-1} K_t^{t-3} \cdots K_t^{s+1}]$ both in the first k rows. This contradiction proves the theorem. The other cases are similar. ▮

Note that the results are best possible in that for $A = [K_t^t K_t^{t-2} \cdots K_t^s]$, $B = [K_t^{t-1} K_t^{t-3} \cdots K_t^{s+1}]$ we have $A^{(s)} = B^{(s)}$, $A^{(s+1)} = B^{(s+1)}, \ldots, A^{(t-1)} = B^{(t-1)}$ as can be verified by some simple counting. Although we have shown that in certain cases $A^{(s)}, A^{(s+1)},..., A^{(t-1)}$ determine a unique A with certain properties, we have not presented an algorithm to find such an A. Only in the case $s=t-1$ of Theorem 2.3 do we have an algorithm, essentially proven for $t=3$ in [16] and easily extended using Quinn's results [13].

Our forbidden configurations are so complex that any configuration is contained in one of them. If F is a configuration contained in a configuration C then any property holding for matrices with no configuration C must hold for matrices with no configuration F. Let F be a configuration of size k by l with no repeated columns. One easily verifies that the first k rows of $[K_{k+1}^{k} K_{k+1}^{k-2} \cdots K_{k+1}^{0}]$ for $k \equiv 0 \pmod 2$ or $[K_{k+1}^{k+1} K_{k+1}^{k-1} \cdots K_{k+1}^{0}]$ for $k \equiv 1 \pmod 2$ has every column on k rows and so contains F. Thus we may obtain the following general configuration theorem as an interesting corollary to Theorem 2.3.

COROLLARY 2.6. *Let F be a configuration of size k by l with no repeated columns. Let A, B be* (0, 1)*-matrices with no configuration F and with $A^{(1)} = B^{(1)}, A^{(2)} = B^{(2)},..., A^{(k)} = B^{(k)}$. Then A and B are the same apart from a column permutation and columns of zeros.*

3. BOUNDS FOR FORBIDDEN CONFIGURATIONS

In this section we apply a configuration theorem, Proposition 2.5, to obtain bounds as described in the Introduction. The starting point is this result due to Ryser, later generalized by Quinn [10].

THEOREM 3.1 [14]. *Let A be an m by n (0, 1)-matrix with no repeated columns, column sums at least 2, and no configuration $[K_3^2]$. Then $n \leqslant \binom{m}{2}$.*

Ryser's proof uses indeterminates in a matrix equation. Let X be the diagonal matrix of order n with entries $x_1, x_2,..., x_n$ on the diagonal where $\{x_1, x_2,..., x_n\}$ are independent indeterminates. Let

$$Y = AXA^{\mathrm{T}}. \tag{3.1}$$

Then the (i, j) entry of Y $(i \neq j)$ is the 2-fold row intersection vector rows i and j of A written with basis $\{x_1, x_2,..., x_n\}$. One is tempted to stop here and not consider k-fold intersections for arbitrary k because there is no matrix equation. After a while, one realizes that the matrix equation is not necessary in the proof and by these means we prove the following result, generalizing results in [2], using the same proof techniques.

THEOREM 3.2. *Let A be an m by n $(0, 1)$-matrix with column sums at least s and with no configuration $[K_t^t K_t^{t-1} \cdots K_t^s]$. Then the maximum number of linear independent vectors (over the rationals) chosen from all k-fold row intersection vectors, for $k = s, s+1, \ldots, t-1$, is equal to the number of distinct columns in A.*

Proof. Consider $A = (a_{ij})$ as indexing m subsets $S_1, S_2, \ldots, S_m$ of an n-set $\{x_1, x_2, \ldots, x_n\}$, where $a_{ij} = 1$ if and only if $x_j \in S_i$. We circumvent the matrix equation (3.1) by defining directly

$$Y_I = \sum \left\{ x_j \,\middle|\, j \in \bigcap_{i \in I} S_i \right\}. \tag{3.2}$$

When $I = \{i_1, i_2, \ldots, i_k\}$, then Y_I corresponds to the k-fold row intersection vector of rows $i_1, i_2, \ldots, i_k$ when Y_I is considered as a vector in Q^n with basis $\{x_1, x_2, \ldots, x_n\}$. Note that $y_\varnothing = x_1 + x_2 + \cdots + x_n$ using the usual definitions.

Repeated columns in A can be deleted without affecting the linear independence of the row intersections. Thus we may assume A has n distinct columns. We immediately deduce that n, the number of distinct columns, is at least the number of linearly independent row intersection vectors since n is the dimension of the space containing them.

Assume n is greater than the number of linearly independent row intersection vectors. Consider the following equations in the variables $x_1, x_2, \ldots, x_n$,

$$\{ Y_I = 0 \mid I \subseteq \{1, 2, \ldots, m\}, s \leqslant |I| \leqslant t-1 \}.$$

Since the number of variables n exceeds the number of linearly independent equations, by assumption, we can find rational and hence integral values $e_1, e_2, \ldots, e_n$, not all zero, for $x_1, x_2, \ldots, x_n$. Since every variable occurs in some equation (either the column sums are at least 1 or for $s = 0$, the equation $y_\varnothing = 0$ verifies the claim), some e_is are positive and some are negative. Define A_1, A_2 as follows. For all i with $e_i > 0$, A_1 contains column i of A repeated e_i times. For all j with $e_j < 0$, A_2 contains column j of A repeated $-e_j$ times. Then it follows that $A_1^{(s)} = A_2^{(s)}, A_1^{(s+1)} = A_2^{(s+1)}, \ldots, A_1^{(t-1)} = A_2^{(t-1)}$. Apply Proposition 2.5 since we have that A_1 and A_2 have no matching columns. Combining the configurations that exist in A_1 and A_2 in the same t rows, we obtain that the forbidden configuration $[K_t^t K_t^{t-1} \cdots K_t^s]$ occurs in A. This contradiction proves the theorem. ∎

COROLLARY 3.3. *Let A be an m by n $(0, 1)$-matrix with no repeated columns, column sums at least s, and with no configuration $[K_t^t K_t^{t-1} \cdots K_t^s]$. Then*

$$n \leqslant \binom{m}{t-1} + \binom{m}{t-2} + \cdots + \binom{m}{s}. \tag{3.3}$$

Proof. The right-hand side is simply the number of k-fold row intersection vectors for $k=s, s+1, ..., t-1$ and so, by Theorem 3.2, it certainly yields a bound. For this forbidden configuration the bound is tight since $[K_m^{t-1}K_m^{t-2}\cdots K_t^s]$ satisfies the hypothesis and yields equality in (3.3). ∎

This bound, in the case $s=0$, was proven by Sauer [18] and Perles and Shelah [19]. Our linear independence argument was obtained independently of the proof of Frankl and Pach [16]. An easy proof of the general bounds has since been provided by Anstee and Murty [15], but the insight provided by Theorems 2.4 and 3.2 is lost. Quinn [13], proved the result for the forbidden configuration K_k^{k-1} with $s=k-1$.

The following result is a direct corollary to Corollary 3.3 using the same observation yielding Corollary 2.6.

THEOREM 3.4. *Let F be a k by l $(0, 1)$-matrix with no repeated columns and column sums at least s. Let A be an m by n $(0, 1)$-matrix with no repeated columns and no configuration F. Then*

$$n \leqslant \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{s}. \tag{3.4}$$

Proof. Simply note that F is contained in $[K_k^k K_k^{k-1} \cdots K_k^s]$ and use Corollary 3.3. In Section 4 we will see that for certain F this bound is tight. ∎

THEOREM 3.5. *Let F be a k by l $(0, 1)$-matrix such that a column of F appears at most t times. Let A be an m by n $(0, 1)$-matrix with no repeated columns. Then*

$$n \leqslant l(t-1)\binom{m}{k} + \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} \tag{3.5}$$

and

$$n \leqslant \binom{m}{s-1} + \binom{m}{s-2} + \cdots + \binom{m}{0} \qquad \textit{for } s = k + \lceil \log_2 t \rceil. \tag{3.6}$$

Proof. The bound (3.5) is given by a pidgeonhole principle. Let F' be the matrix obtained by deleting from F any repeated columns. We claim that in a matrix A with no repeated columns of size n by $l(t-1)\binom{m}{k} + \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} + 1$ that the configuration F' can be found to occur t times, each time in a disjoint set of columns, in the same k rows. Thus A would contain the configuration F. Note that after $\binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} + 1$ columns, we must find a copy of F' by Theorem 3.4. Delete the columns containing the copy of F' from consideration and add enough columns (l will do) so that we again have at least $\binom{m}{k-1}+\binom{m}{k-2}+\cdots+\binom{m}{0}+1$ distinct columns. Then we can find another copy of F'. Thus in A we can find $(t-1)\binom{m}{k}+1$ copies of F' each in a separate set of columns. By the pidgeonhole principle, at least t will occur in the same set of k rows. This proves the claim and hence the bound.

The second bound is weaker but it is much easier to visualize. Simply take F and append enough rows to it so that the repeated columns are distinguished. If we add $\lceil \log_2 t \rceil$ rows, then we can extend a column in $2^{\lceil \log_2 t \rceil}$ ways and so distinguish $2^{\lceil \log_2 t \rceil} > t$ columns of F. The result now follows by applying Theorem 3.4. Bill Cook suggested this approach in view of Theorem 3.4. ∎

The first bound is nearly independent of l, a most surprising result. We do not have examples to see if the bound is very good but it is reassuring that it is so close to Theorem 3.4. We have omitted restrictions on column sums since such precision does not seem warranted.

4. Extremal Matrices

Having presented the bound of Theorem 3.4 for a configuration F of size k by l with no repeated columns, one is left with the question of whether it is a good bound. If F has a column of 1s then $[K_t^{t-1}K_t^{t-2}\cdots K_t^s]$ verifies that the bound is tight. A much more interesting class of examples is provided by the result of Füredi and Quinn [11].

Proposition 4.1. *There exist m by $\binom{m}{t-1}$ (0, 1)-matrices with no repeated columns, column sums at least $t-1$, and no configuration $[K_t^{t-1}]$.*

Quinn proved much more concerning the structure of such matrices, generalizing the result in [1]. Ryser proved this result for $t=3$ [17]. Let A be a matrix as described in Proposition 4.1. If F contains the configuration $[K_t^{t-1}]$, then the matrix $B=[K_t^sK_t^{s+1}\cdots K_t^{t-2}A]$ meets the bound of Theorem 3.4. In the case $s=0$ we can take the (0, 1)-complement of B with the resulting column of 0s deleted to obtain a matrix with no $[K_t^1]$ which meets the bounds of Theorem 3.4. Finally we note that $[K_m^mK_m^{m-1}\cdots K_m^{m-t+1}]$ is a matrix with no $[K_t^0]$ meeting the bounds of Theorem 3.4. Thus if either F contains the configuration $[K_t^1]$ or $[K_t^0]$, then the bound of Theorem 3.4 is tight. Thus we have a large class of Fs for which the bound is tight.

The above constructions yield all the matrices meeting the bound when the forbidden configuration is in the list $[K_1^1]$, $[K_1^0]$, $[K_2^2K_2^0]$, $[K_2^2]$, $[K_2^0]$, $[K_2^1]$, $[K_3^3K_3^1]$, and $[K_3^2K_3^0]$ (the results for the last two appear in [2]). Beyond this it seems that other extremal matrices may arise. For example, we could delete a column of column sum 2 from $[K_4^4 K_4^3 K_4^2 K_4^1 K_4^0]$ to obtain a $m \times (\binom{m}{3} + \binom{m}{2} + \binom{m}{1} + \binom{m}{0})$ matrix with no configuration $[K_4^4 K_4^2 K_4^0]$. Some general results do hold for extremal matrices.

THEOREM 4.2. *Let A be an m by $(\binom{m}{t-1} + \binom{m}{t-2} + \cdots + \binom{m}{s})$ $(0, 1)$-matrix with no repeated columns, column sums at least s and with no configuration $[K_t^t K_t^{t-1} \cdots K_t^s]$. Then there exists a permutation matrix P of order $\binom{m}{t-1} + \binom{m}{t-2} + \cdots + \binom{m}{s}$ such that $A \geqslant [K_m^{t-1} K_m^{t-2} \cdots K_m^s]\, P$.*

Proof. This is the usual elementwise $\geqslant$ for matrices. Associate with each column i of A the set S_i,

$$S_i = \{I \subseteq \{1, 2,..., m\} \mid s \leqslant |I| \leqslant t-1,$$
$$\text{column } i \text{ of } A \text{ has 1s in rows indexed by } I\}. \tag{4.1}$$

Then P corresponds to an SDR of this set system. Let $J \subseteq \{1, 2,..., \binom{m}{t-1} + \binom{m}{t-2} + \cdots + \binom{m}{s}\}$. Let B be the m by r submatrix of A consisting of all columns of A indexed by J. Since A does not have the forbidden configuration then neither does B. We find that $|\bigcup\{S_i | i \in J\}|$, the number of nonzero k-fold row intersection vectors is at least the number of linearly independent row intersection vectors which is the number of columns of B by Theorem 3.2. Thus for any J,

$$|\bigcup\{S_i | i \in J\}| \geqslant |J|. \tag{4.2}$$

We deduce, by P. Hall's theorem, that the set system has an SDR as desired [15]. ▮

We may interpret this result as saying that $[K_m^{t-1} K_m^{t-2} \cdots K_m^s]$ acts as a skeleton for these extremal matrices to which additional 1s may be added. This result clearly holds when we simply consider a configuration contained in $[K_t^t K_t^{t-1} \cdots K_t^s]$. Theorem 4.2 was proved with $t=3$, $s=0$ using the forbidden configurations $[K_3^3 K_3^1]$ or $[K_3^2 K_3^0]$ and this proof mimics the proof there [2]. In the case $t=3$, $s=2$ with the forbidden configuration K_3^2, the matrix P was shown to be unique [1]. Quinn generalized this result to the case $t=k$, $s=k-1$ and the forbidden configuration K_k^{k-1} [13].

An inductive buildup of these extremal matrices is suggested by the following observation.

THEOREM 4.3. *Let A be an m by $(\binom{m}{t-1} + \binom{m}{t-2} + \cdots + \binom{m}{s})$ $(0, 1)$-matrix with no repeated columns, column sums at least s and with no configuration $[K_t^t K_t^{t-1} \cdots K_t^s]$. Then any submatrix B obtained by taking l rows of A and deleting repeated columns and columns of sum less than s is of size l by $(\binom{l}{t-1} + \binom{l}{t-2} + \cdots + \binom{l}{s})$.*

Proof. Note that in A, since it meets the bound of Corollary 3.3, all the k-fold row intersection vectors ($k = s, s+1, \ldots, t-1$) are linearly independent and so they are all linearly independent for B for $k = s, s+1, \ldots, \min(l, t-1)$. Deleting columns with column sum less than s or repeated columns does not affect this. Applying Theorem 3.2, we obtain the result. ∎

Again, this result was noted in [1, 2] and by Quinn [13], for certain special forbidden configurations. An alternative proof can be found in [5].

5. Bounds for Forbidden Submatrices

It turns out that handling forbidden submatrices is not much more difficult than handling forbidden configurations. We say that a configuration C *always has* a submatrix F if every representative of C has F as a submatrix. Thus if A is a matrix with no submatrix F, then A has no configuration C.

Thus to prove a version of Theorem 3.4 for forbidden submatrices, we need to find for each submatrix F, a configuration C which always has F as a submatrix. In view of Corollary 3.3, we need only consider the t by 2^t configuration

$$P_t = \left[K_t^t K_t^{t-1} \cdots K_t^0 \right], \tag{5.1}$$

all columns on t rows. Note that any representative of P_t has the same columns, perhaps in a different order.

LEMMA 5.1. *The configuration P_t always has any k by l $(0, 1)$-matrix F as a submatrix for $t \geqslant kl$ and for $t \geqslant 13k \log_2 l$.*

Proof. The first bound is proven by induction on l. The result is true for $l = 1$ easily. Consider a representative of P_{kl} for some $l > 1$. Say it has l copies of P_k in rows $1, 2, \ldots, l$, where all the columns of the $(i+1)$th copy of P_k occur to the right of all the columns of the ith copy of P_k. Then the ith copy of P_k certainly has the ith column of F and so this representative of P_{kl} has F as a submatrix.

If this does not occur, then there are two columns α, β of P_k such that all the $2^{k(l-1)}$ columns of P_{kl} with α in rows $1, 2, \ldots, k$ occur to the left of at least $2^{k(l-1)} - (l-2)$ columns of P_{kl} with β in rows $1, 2, \ldots, k$. Thus in rows $k+1, k+2, \ldots, kl$ we have a copy of $P_{k(l-1)}$ occurring to the left of a copy of $P_{k(l-1)}$ minus up to $l-2$ columns. By induction, $P_{k(l-1)}$ has a submatrix $\bar{F}$ which is F minus its last column, say ω. We need only verify that any representative of $P_{k(l-1)}$ with up to $l-2$ columns deleted has a column with ω as a submatrix in k selected rows for any such selection. But for a given choice of k rows, there are $2^{k(l-2)}$ such columns in $P_{k(l-1)}$. Since $2^{k(l-2)} > l-2$ for $k \geqslant 1$, $l \geqslant 2$, such columns exist. Thus every representative of P_{kl} has F as a submatrix as desired.

The second bound is a pidgeonhole principle analogous to the argument in Theorem 3.5. Let F' be the configuration $[K_k^{s_1} K_k^{s_2} \cdots K_k^{s_p}]$ obtained from F, where $s_1, s_2, ..., s_p$ are the distinct column sums of F. In any $\binom{t}{k-1} + \binom{t}{k-2} + \cdots + \binom{t}{0} + 1$ distinct columns of P_t, one will find a configuration F' by Theorem 3.4. Thus after $(l-1)\binom{t}{k} + 1$ such sets of columns in P_t we will have $(l-1)\binom{t}{k} + 1$ copies of F' each copy entirely to the right of the preceding copy. But by the pidgeonhole principle, there will be l copies in some set of k rows. From these l copies one can extract F as a submatrix. Thus P_t will have F as a submatrix if

$$\left((l-1)\binom{t}{k}+1\right)\left(\binom{t}{k-1}+\binom{t}{k-2}+\cdots+\binom{t}{0}+1\right) \leqslant 2^t, \tag{5.2}$$

i.e., there are enough columns in P_t. We solve for a t which satisfies (5.2) by using various approximations. We have

$$\left((l-1)\binom{t}{k}+1\right)\left(\binom{t}{k-1}+\binom{t}{k-2}+\cdots+\binom{t}{0}+1\right) \leqslant lk\binom{t}{k}^2. \tag{5.3}$$

Now, using Stirling's formula, we have

$$lk\binom{t}{k}^2 \leqslant lk\left(\frac{t^k}{k!}\right)^2 < lk\left(\frac{t^k}{\sqrt{2\pi k}\, k^k\, e^{-k}}\right)^2 = \frac{l}{2\pi}\left(\frac{te}{k}\right)^{2k}. \tag{5.4}$$

Writing this as a power of 2, we have

$$\frac{l}{2\pi}\left(\frac{te}{k}\right)^{2k} = 2^{\log_2 l/2\pi + 2k \log_2(te/k)}. \tag{5.5}$$

For $t = 13k \log_2 l$, we find that

$$\log_2 l/2\pi + 2k \log_2(te/k) < t, \tag{5.6}$$

and so (5.2) holds. This is clearly not the best possible, especially for given k and l. However the form of the answer is interesting. Bruce Richmond provided some helpful comments concerning the determination of t. ∎

THEOREM 5.2. *Let A be an m by n $(0, 1)$-matrix with no repeated columns. Let F be a $(0, 1)$-matrix of size k by l and assume A does not have F as a submatrix. Then*

$$n \leqslant \binom{m}{t-1} + \binom{m}{t-2} + \cdots + \binom{m}{0}, \tag{5.7}$$

where $t = kl$ or $t = 13k \log_2 l$.

Proof. Simply note that P_t always has F as a submatrix for the given values of t by Lemma 5.1. Apply Corollary 3.3. ∎

It is easy to show that the bounds of Lemma 5.1 are not best possible for certain F and so surely Theorem 5.2 is not best possible. Nonetheless, the bounds of Theorem 5.2 are surprisingly good, being close to those of Theorem 3.4 which are best possible in certain cases. After all, forbidding a submatrix is potentially much weaker than forbidding the configuration of which it is a representative.

The following forbidden submatrix theorem follows from some results on totally balanced matrices (see the Introduction). Totally balanced matrices can be characterized as precisely those matrices for which there is a row and column permutation which has no submatrix

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \tag{5.8}$$

This result was proved by Brouwer and Kolen [7] and Anstee and Farber [4]. This forbidden submatrix property is vital in any algorithmic work.

PROPOSITION 5.3. *Let A be an m by n $(0, 1)$-matrix with no submatrix* (5.8). *Then $n \leqslant \binom{m}{2} + \binom{m}{1} + \binom{m}{0}$ and this bound is best possible.*

Proof. We note that K_3^2 always has (5.8) as a submatrix and so the bound holds. Matrices meeting the bound with no configuration K_3^2 are constructed in [1] and shown to be totally balanced. The above results complete the proof. ∎

Thus we have beaten the bound of Lemma 5.1 using the matrix (5.8). As a final remark, note that even if you can get best possible results for the t of Lemma 5.1, there is no reason to expect that the bound you get for Theorem 5.2 is best possible.

REFERENCES

1. R. P. ANSTEE, Properties of (0, 1)-matrices with no triangles, *J. Combin. Theory Ser. A* **29** (1980), 186–198.
2. R. P. ANSTEE, Properties of (0, 1)-matrices without certain configurations, *J. Combin. Theory Ser. A* **31** (1981), 256–269.
3. R. P. ANSTEE, Extensions of the notion of conformality in hypergraphs, Proceedings of the 14th Southeastern Conf. on Combinatorics, Graph Theory, and Computing, *Congressus Numerantium* **39** (1983), 83–88.
4. R. P. ANSTEE AND M. FARBER, Characterizations of totally balanced matrices, *J. Algorithms* **5** (1984), 215–230.
5. R. P. ANSTEE AND U. S. R. MURTY, Matrices with forbidden subconfigurations, *Discrete Math.* **54** (1985), 113–116.

6. C. BERGE, "Graphs and Hypergraphs," North-Holland, Amsterdam, 1973.
7. A. E. BROUWER AND A. KOLEN, "A Super-balanced Hypergraph has a Nest Point," Report ZW 146/80, Mathematisch Centrum, Amsterdam, 1980.
8. A. BRUEN AND R. SILVERMAN, Switching sets in $PG(3, q)$, *Proc. Amer. Math. Soc.* **43** (1974), 176–180.
9. D. DE CAEN, "On Turan's Hypergraph Problem," Ph.D. thesis, University of Toronto, July 1982.
10. P. FRANKL AND J. PACH, On the number of sets in a null t-design, *European J. Combin.* **4** (1983).
11. Z. FÜREDI AND F. QUINN, Traces of finite sets, *Ars Combinatoria*, in press.
12. L. LOVÁSZ, "Combinatorial Problems and Exercises," North-Holland, Amsterdam, 1979.
13. F. QUINN, (0, 1)-matrices containing no permutation matrices of specified order, preprint.
14. H. J. RYSER, The term rank of a matrix, *Canad. J. Math.* **10** (1958), 57–65.
15. H. J. RYSER, Combinatorial configurations, *Siam. J. Appl. Math.* **17** (1969), 593–602.
16. H. J. RYSER, Intersection properties of finite sets, *J. Combin. Theory Ser. A* **14** (1973), 79–92.
17. H. J. RYSER, A fundamental matrix equation for finite sets, *Proc. Amer. Math. Soc.* **34** (1972), 332–336.
18. N. SAUER, On the density of families of sets, *J. Combin. Theory Ser. A* **13** (1972), 145–147.
19. S. SHELAH, A combinatorial problem; Stability and order for models and theories in infinitary languages, *Pacific J. Math.* **41** (1972), 247–261.